Scalable metadata environments (MDE): artistically-impelled immersive environments for large-scale data exploration

Ruth G. West*[a], Todd Margolis[b], Andrew Prudhomme[b], Jurgen P. Schulze[b], Iman Mostafavi[d]
JP Lewis[c] , Joachim Gossmann[b], Rajvikram Singh[b],
[a]University of North Texas, 1155 Union Circle, Denton TX 76203; [b]UCSD 9500 Gilman Dr., La Jolla, CA 92093; [c]Victoria University Wellington, NZ; [d]Limbic Software, San Francisco, CA 94301

ABSTRACT

Scalable Metadata Environments (MDEs) are an artistic approach for designing immersive environments for large scale data exploration in which users interact with data by forming multiscale patterns that they alternatively disrupt and reform. Developed and prototyped as part of an art-science research collaboration, we define an MDE as a 4D virtual environment structured by quantitative and qualitative metadata describing multidimensional data collections. Entire data sets (e.g.10s of millions of records) can be visualized and sonified at multiple scales and at different levels of detail so they can be explored interactively in real-time within MDEs. They are designed to reflect similarities and differences in the underlying data or metadata such that patterns can be visually/aurally sorted in an exploratory fashion by an observer who is not familiar with the details of the mapping from data to visual, auditory or dynamic attributes. While many approaches for visual and auditory data mining exist, MDEs are distinct in that they utilize qualitative and quantitative data and metadata to construct multiple interrelated conceptual coordinate systems. These "regions" function as conceptual lattices for scalable auditory and visual representations within virtual environments computationally driven by multi-GPU CUDA-enabled fluid dyamics systems.

Keywords: **Keywords:** immersive multiscale, multiresolution visualization, art-science collaboration, spatialized multi-channel interactive audio, audio rendering, audio spatialization

1. MAKING THE ABSTRACT EXPERIENTIAL

As we race towards a “digital universe” of 40 trillion gigabytes by 2020 that encompasses the full scope of human endeavor from science to the economy, humanities, telecommunication and the arts, we are challenged not only by its size, but its ephemerality[1]. We must also come to terms with its incompleteness and our inability to effectively search, aggregate and cross-reference its myriad elements[2]. While data is often considered a resource, a raw material that can be manipulated and refined along a continuum from information-to-knowledge-to-wisdom[3] fundamentally there is, and may always be, a gap between the data, the underlying phenomena it represents, and the meaning ascribed to it. One can devise rules to assign meaning to the output of rule-based systems, yet the output itself must be interpreted in turn, leading to an infinite regress[4]. Generating, storing, accessing, representing and interpreting data also necessarily involve subjective choices. This is not always acknowledged nor made explicit. Through choices such as what to sample, the sampling resolution, file formats, what gets discarded versus stored when the data is too large to retain all of it, or the database schemas utilized in managing it, unspoken framing narratives arise that encode agreed upon assumptions about what the creators think they will find in the data, what they think they can know. Framing narratives also arise from our choice of representational schemas, statistics, and algorithms, displays, interaction technologies, and metaphors. Recording complex phenomena from the personal to the global as digital data with technologies that often transcend the capacities of our senses (E.g. fitness wearables, terrestrial observatories, ultra-high resolution sub-cellular imaging, databases of consumer transactions, genomics etc.) creates digital repositories with information content rich enough to produce an enormous number of observations. Yet, an individual or given domain expert can only generate a limited number of interpretations, mostly guided by their specific expertise and the respective framing narratives of data creation and representation. These observations combined with the emergence of ArtScience as an approach for creating new ways of seeing and knowing through hybrid strategies[5] motivate our pursuit of aesthetic and artistically-impelled approaches to support intuitive exploration of large data collections that transcend disciplinary boundaries and individual

The Engineering Reality of Virtual Reality 2014, edited by Margaret Dolinsky, Ian E. McDowall,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 9012, 901205 · © 2014 SPIE-IS&T
CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2038673

expertise. Many existing visualization techniques seek to preserve quantitative transparency in the data display[6]. For example, direct visualizations present a one-to-one mapping of data attributes to visual elements. This paper presents our exploration of creative practice based methods for working with large and highly dimensional data that do not focus on direct visualization/sonificaiton and one-to-one mappings because the number of potential mappings in abstract and highly dimensional data is vast. In fact, it is more than combinatorial, since it is the number of possible mapping programs, and as such systematically exploring this space using an engineering or optimization inspired approach is likely intractable. In the sections below we describe our prototype design for metadata environments, one approach for creating aesthetic and experiential encounters with vast and abstract data spanning scales from nature to culture.

2. METADATA ENVIRONMENTS

We define a scalable metadata environment (MDE) as a virtual space partitioned in to regions based on metadata relevant to one or more data collections. Regions function as conceptual lattices for dynamic and scalable visual and auditory representations. They facilitate embodied exploration in a manner akin to scaffolded environments in which each sub-region establishes distributed patterns that contribute to a larger pattern-structure that humans can simultaneously engage and co-create[7]. In parallel to the way that an architectural space has sub-spaces reflecting human intention and externalized memory or guided cognition, metadata environment regions collectively represent the "space" and "pattern" containing a data set existing at immaterial scales and make it available for embodied exploration.

To provide the data framework for ATLAS *in silico* (http://www.atlasinsilico.net), an interactive artwork blending new media art, electronic music, virtual reality, data visualization and sonification, with metagenomics[8] we developed a prototype MDE for the Global Ocean Sampling (GOS) expedition dataset. Creating this artwork also included developing schemas for scalable visual[9] and auditory[10] data representation, along with novel hybrid strategies for 10.1 multi-channel interactive audio spatialization and localization[11]. These elements were integrated with infrared head and hand motion tracking for enabling user interaction within the immersive environment. The GOS (2003 - 2006) conducted by the J. Craig Venter Institute, studies the genetics of communities of marine microorganisms throughout the worlds oceans. It produced a vast metagenomics dataset with "far-reaching implications for biological energy production, bioremediation, and creating solutions for reduction/management of greenhouse gas levels in our biosphere [12]." The data contains millions of DNA sequences and their associated predicted amino acid (protein) sequences. These predicted sequences, called "ORFs" (Open Reading Frames), candidates for putative proteins, are subsequently validated by a variety of bioinformatics analyses. It also includes a series of metadata descriptors, such as temperature, salinity, depth of the ocean, and depth of the sample, latitude and longitude of the sample location that describe the entire GOS data collection. For ATLAS *in silico* we utilized the entire first release of the GOS which contained 17.4 million ORFs [*ibid*]. Analysis of the GOS is ongoing and the dataset, available online via the CAMERA portal, is comprised of over 60 million sequences[13].

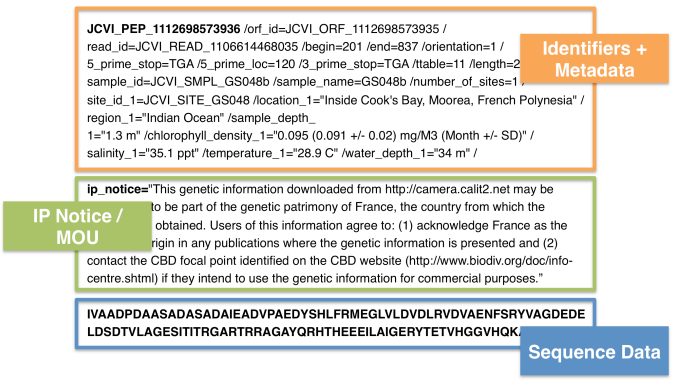

Fig. 1. Three major components of a GOS record: identifiers and metadata, IP notice resulting from MOU and sequence data.

Each database record within the GOS spans scales from the imperceptible informational scales of genetic sequences to palpable environmental metrics including water temperature or salinity, to IP notices generated by country-specific

MOUs along the route of the circumnavigation. The later can be indexed to macro socio-economic variables at global scales such as CO2 emissions, internet usage per capital or infant mortality rates. In this way each record of the GOS spans both nature and culture.

The MDE constructed for the GOS enables exploration of the 17.4 M GOS ORFs (100% of the first GOS release[12]) in a single virtual environment in which the entire data set is visualized and sonified simultaneously at interactive frame rates. Each region within the environment is constructed from one metadata category with its value range mapped to the depth (z) axis, along with contextual and conceptual dimensions such as GPS location, mapped to the other (x, y) axes. The entire MDE environment is instantiated within a fluid force computed as described in[14] wherein each particle in the fluid dynamics simulation corresponds to a single database record. The movement of particles within this system reveals the specific values of all metadata descriptors for each record. The concept of the "metadata cell" or sub-region of the virtual environment is integral to the design of MDEs (see figure 2). Each metadata cell represents specific attributes of the entire data collection, with each region representing all possible values of each metadata category. This concept is central to the mechanisms underlying the dynamic sifting/sorting that enables emergent patterns to develop revealing structures within the entire data set influenced by the fluid forces within the virtual environment. Data (particles) are placed in an MDE within the fluid simulation at random starting positions. Since each sub-region is essentially a volume with an individual coordinate system, the overall environment can be seen as constructed by a large coordinate lattice. As particles (data elements) enter and move about regions their movement and interactions are constrained by the metadata properties for each region as well as by the metadata annotations that each data record (particle) carries. Over time as the data moves throughout the space an overall pattern emerges. The patterns result from kinesthetic movement of clusters of data records moving together in space.

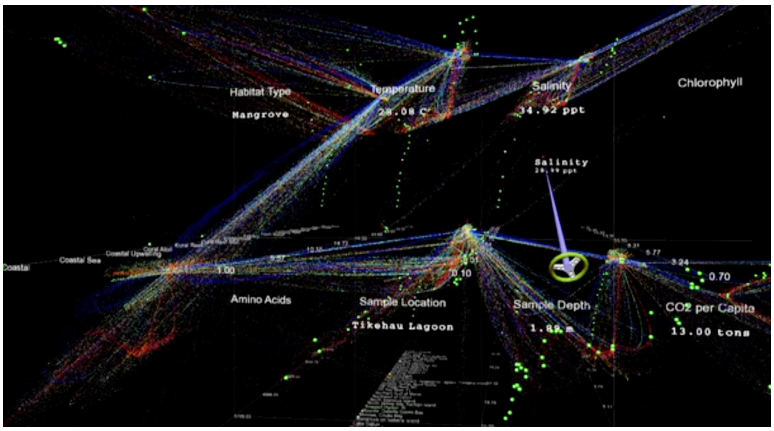

Fig. 2. Metadata cells (regions) and visual, behavioral and auditory encodings for the GOS MDE.

2.1 Evolving the MDE framework

Over the course of its development the MDE framework for the GOS has evolved from a single threaded CPU implementation providing interaction with 20,000 records, to a single-GPU based CUDA implementation allowing for exploration of 1M records to a multi-GPU and CUDA based parallelized and partitioned framework enabling interaction with 17.4M records comprising the entire first release of the GOS. Figure 3 shows an overview and detail inset of the MDE for the GOS.

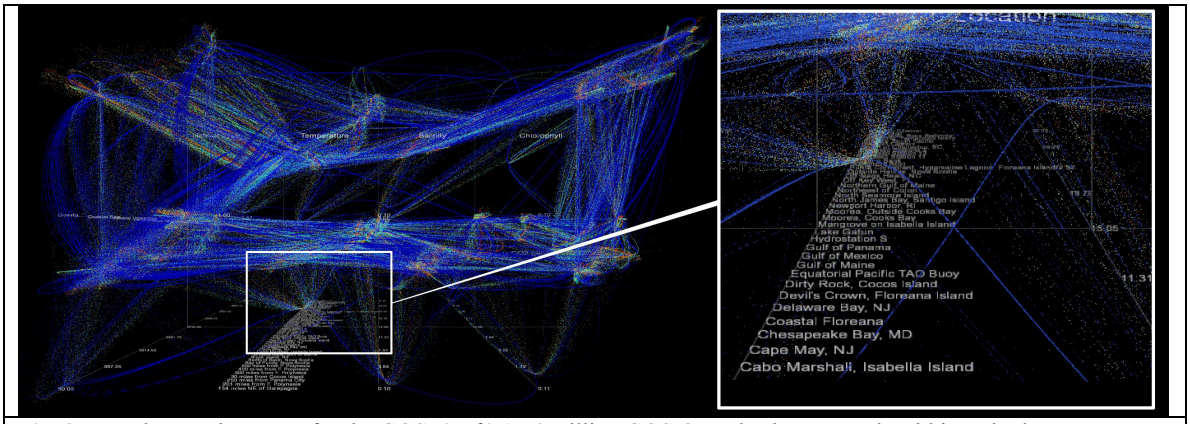

Fig. 3. Metadata environment for the GOS: (Left) 17.4 million GOS ORF database records within a single MDE. Detail: (Inset/Right) GOS ORF records (particles) cluster in to streamline-like spatio-temporal patterns when records share metadata characteristics. Records with differing sets of characteristics move in distinct trajectories creating distributed patterns. Cool to warm pseudocolor map indicates the number of sampling sites each ORF assembles across. Blue = 1 site, deepest red = 24 sites. (Right)

The partitioned and parallelized structure of the framework supports filtering operations to create subsets of ~ 1M records directly from the entire 17.4M GOS record dataset. The filtered subset of records is immediately explored within the MDE while the system simultaneously continues the simulation for the entire dataset thus maintaining the relationship of the subset to the whole. Figure 4, below, summarizes the multiple approaches undertaken.

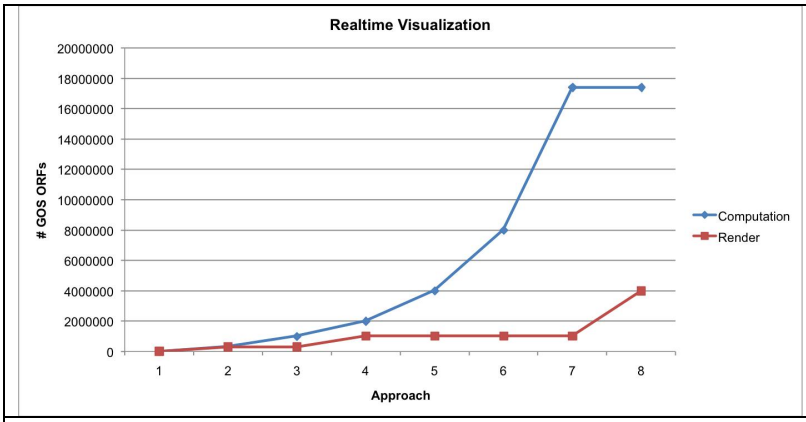

Fig. 4. Multiple approaches for realtime visualization with frame rates ranging from 12 to 39 FPS were achieved with multiple combinations of GPUs and partitioning approaches.
Approaches legend:
(1) single-threaded, CPU-based simulation (non-VBO) (non-CUDA);
(2) multi-threaded, CPU-based simulation (non-VBO) (non-CUDA);
(3) multi-threaded, CUDA-enabled simulation (non-VBO) (1 card sim & 1 gfx card render);
(4) multi-threaded, CUDA-enabled simulation (VBO) (1 card sim & 1 gfx card render);
(5) multi-threaded, CUDA-enabled simulation (VBO) (1 card for both sim and render);
(6) multi-threaded, CUDA-enabled simulation on 2 cards 2 GPUs (VBO) & render on another card (Synchronous update);
(7) multi-threaded, CUDA-enabled simulation on 2 cards 3 GPUs (VBO) & render on another card (Synchronous update);
(8) multi-threaded, CUDA-enabled simulation on 2 cards 3 GPUs (VBO & display list, progressive updating) & render on another card

The multi-GPU based CUDA enabled implementation utilizes both parallelization and partitioning strategies to provide user interaction (head and hand tracking) at interactive frame rates along with the underlying simulation of the entire data set and real-time rendering.

Figure 5 demonstrates the result of a filtering operation to select GOS ORFs that assemble in one or more sets of sequences from sampling sites in addition to the collection corresponding to their physical sampling location.

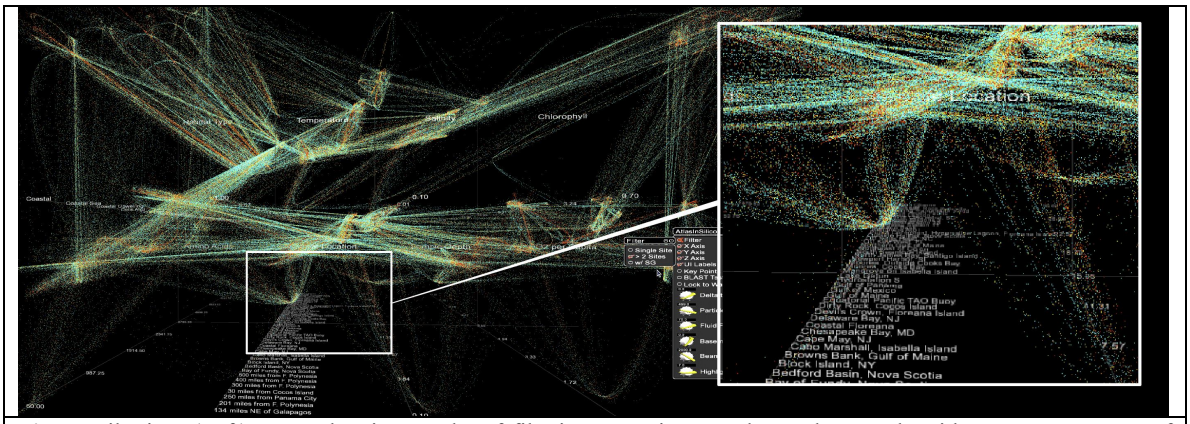

Fig. 5. Filtering: (Left) MDE showing results of filtering operation to select only records with two or more sets of metadata, and therefore sequences identified at two or more sample sites. Detail(Inset/Right) Cool to warm pseudocolor map indicates number of sampling sites. Compare inset to figure 1 inset. Blue = 1 site, deepest red = 24 sites.

In-world user interface features enable drill down/through operations, subset filtering (creating subsets of 1M records), small sub-set selection of 20 records for comparative exploration, and single record selection for detailed exploration. These operations are accompanied by distortion-less detail-in-context viewing via the form of a "traceback" metadata signature presented in the context of the entire data set.

2.2 Scalable Representations and User Interaction

The 4D nature of MDEs creates dynamic patterns from entire data collections. In-world user interface features allow users to interact with and explore data by disrupting these patterns and observing them as they are dynamically reformed. An in-world menu panel (Figure 6 below) enables easy toggling of UI elements and control of simulation parameters. The design and user experience includes several interrelated representations and interactive modalities spanning multiple levels of scale.

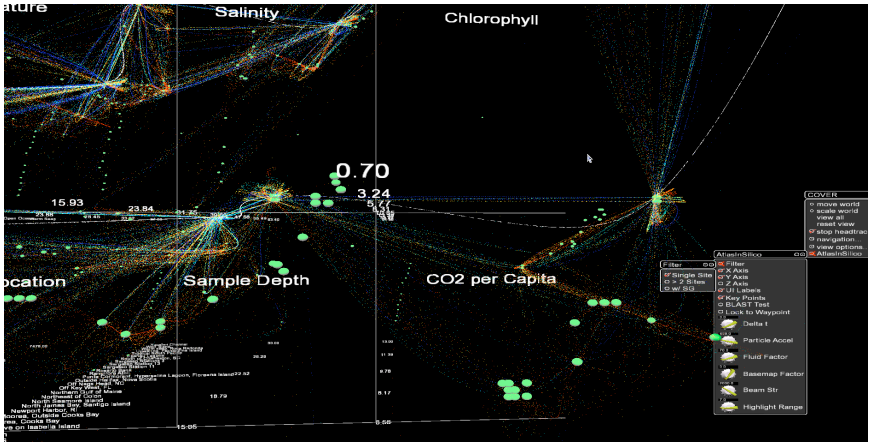

Fig. 6. GOS MDE with in-world control panel

MDEs are designed to reflect similarities and differences in the underlying data or metadata such that patterns can be visually/aurally sorted in an exploratory fashion by an observer who is not familiar with the details of the mapping from data to visual, auditory or dynamic attributes. This requires an approach that generates as wide a range of distinctive patterns as possible. We developed and utilize scalable auditory data signatures (SADS) for data sonicafication[10] and meta-shape grammars[9] in which rule generation and interpretation is mapped to higher-order functions as described in[15] for data visualization as n-dimensional glyphs, and we generate spatio-temporal signatures (which we term "tracebacks") to visualize metadata characteristics at multiple scales. Each record, in addition to being represented by a particle in the MDE fluid force is visually encoded according to metadata annotation sets and behaviorally encoded within the MDE according to both values within the record and metadata values in the context of fluid forces. MDE regions are not only spatially distinct but aurally differentiated so that user interaction with one region elicits distinct auditory features from interaction with other regions. This auditory distinctiveness is generated by differences in the metadata characteristics of each region as well as differences in the properties of the data nearby where a user is exploring the MDE by disrupting the emergent patterns.

The user experience starts with an overview of the entire dataset in the MDE "particle" mode. To go to the next level of scale, users can filter the data to view larger subsets (~1M records) or select a small subset of records nearby a point in the metadata space. The selection places the individual records in the foreground within the context of the entire dataset. This change is represented by changes in visual auditory and behavioral encodings. The entire dataset is placed in to the "background" as the particles (records) lose their color encoding and de-saturate to a grayscale value as the fluid simulation "freezes." Simultaneously the subset of records retain some aspects of the visual encoding and evolve in to distinct glyph structures that also incorporate distinct auditory (SADS) and behavioral encoding. Selecting one record out of this subset transitions to the next level of scale in which an individual record is foregrounded against the entire dataset, while region markers and information from within the specific record emerge as visuals, audio and text. The relationship of the individual record to the entire MDE is revealed by a spatio-temporal signature that incorporates all of the metadata values for the record within the context of each region and the MDE as a whole. Deselecting this record returns it in to the data set, transitioning the view back to the active fluid simulation state in which patterns can be explored, disrupted, reformed, and filtering, selection and drill down/up operations performed.

Figure 7 shows the activity of the rendering pipeline for the meta-shape grammar generating n-dimensional glyphs in-world and in real-time from user selected subsets of 20 records for visual and auditory data exploration. The right-most panel shows the result of selecting a single database record from the set of 20 for visual/auditory exploration. The selected glyph (a shape grammar object visualizing the ORF sequence, biophysiochemical features and metadata) is presented in the context of the spatiotemporal metadata signature (traceback) and both are in the context of target data. The data underlying both the glyph and spatio-temporal signature are sonified.

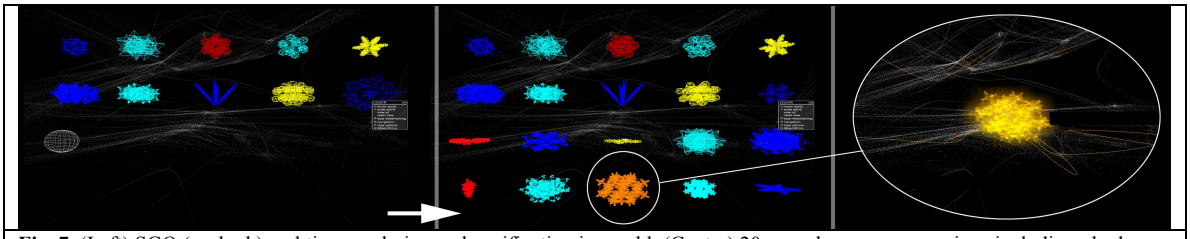

Fig. 7. (Left) SGO (n-glyph) real time rendering and sonification in-world; (Center) 20 records group comparison including glyphs, interaction and sonification; (Right) single record selection, glyph, interaction, sonification, detail-in-context viewing with backgrounded data set and metadata signature traceback.

For the GOS data, in order to combine sequence and geospatial data spaces we established "waypoint" markers within the MDE. This enabled us to add a layer of geo-coded information for each metadata region. Waypoint markers (Figure 8)cross-reference sampling site latitude and longitude with metadata values within each region. Selecting and highlighting waypoints activates a coordinated view to display metadata values across regions and within a region, as shown in figure 8 below.

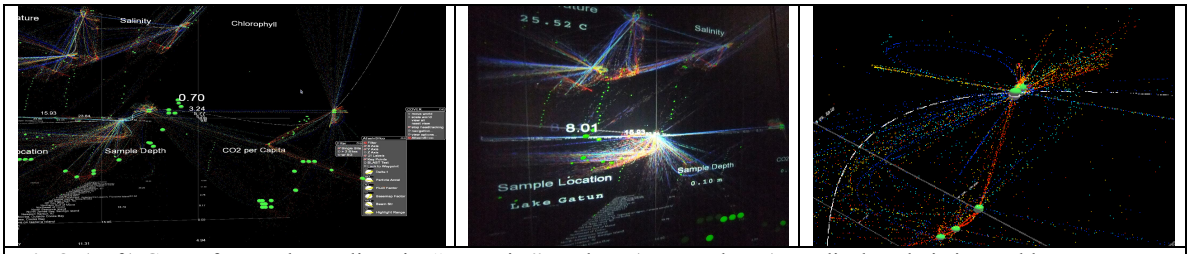

Fig 8. (Left) Geo-referenced sampling site "waypoint" markers (green spheres) are displayed via in-world UI menu. (Center) Browsing the data by selecting a waypoint presents the corresponding metadata coordinated across all regions. Metadata values for sample location, sample depth, temperature and salinity are displayed for a selected waypoint. (Right) GOS ORFs moving in patterns through sample site waypoints.

In addition to the regular grid structure (as in figures 2, 3, 5, 6 and 8), our design for the MDE includes functionality for metadata cells/region juxtaposition to be reconfigured, providing alternative views of spatial temporal patterns within the overall data set in addition to the multiple perspectives available from navigating the initial configuration.

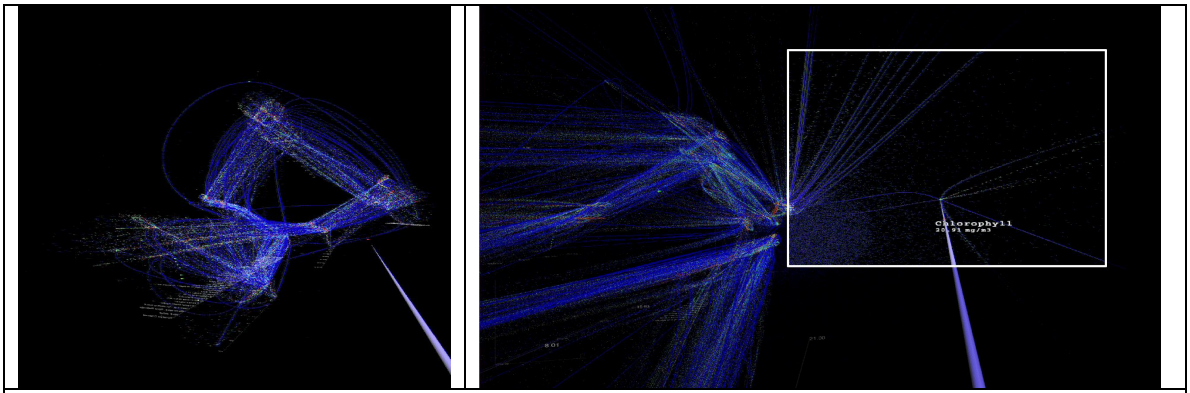

Fig. 9. Users can reconfigure the position of regions to explore alternate juxtapositions of metadata. (Left) 1M records in a user-determined region order and position. (Right) Data explored in the context of chlorophyll values.

2.3 Use case: MDEs as New Views of BLAST

BLAST, the Basic Local Alignment Search Tool[16,17], allows researchers to compare DNA or protein sequences of unknown identity, function, and structure with "knowns" from validated databases, providing a measure of functional or evolutionary similarity or homology among sequences. It is one of the foundational bioinformatics techniques utilized world wide.

Working with CAMERA researchers we developed a use case demonstrating the potential application MDEs in providing novel views of BLAST results that incorporate metadata as a context to BLAST analyses. A BLAST query was run with an individual sequence from the GOS. The small set of "top hit" results for the BLAST query is visualized within the GOS MDE as shown in figures 10 and 11. The multiple types of visual, auditory and behavioral encoding allow users to explore the metadata characteristics of the BLAST results in relation to the target dataset's metadata characteristics within a single MDE. In web-based BLAST user interfaces, users receive tabular lists of metadata attributes that correspond to the records returned by the algorithm. This metadata is analyzed separately and not in the context of either the query or the target dataset as is possible in an MDE.

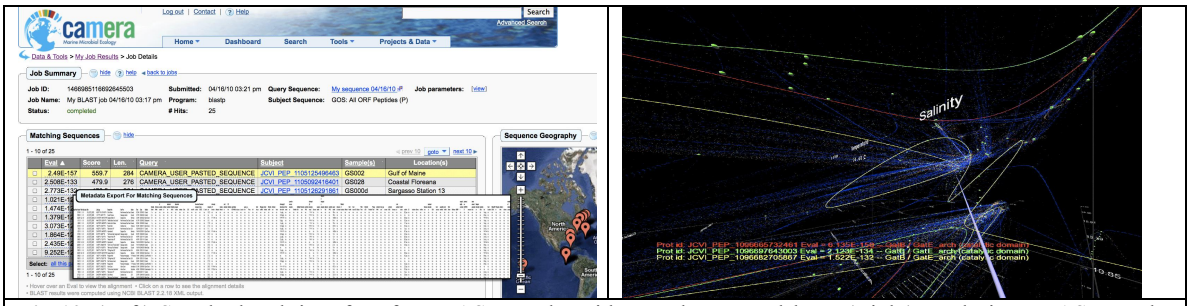

Fig 10. (Left) Standard web interface for BLAST results with metadata spreadsheet. (Right) Exploring BLAST results in an MDE with metadata and target data contexts.

In a standard analysis and user interface, the differences in metadata values for each of the top three "hits" (results) of the blast query would be difficult to see, and even more difficult to see in relation to each other and the overall metadata characteristics of the target database. . Figure 11 below demonstrates the potential for MDEs to augment BLAST analysis by presenting query results in their metadata context.

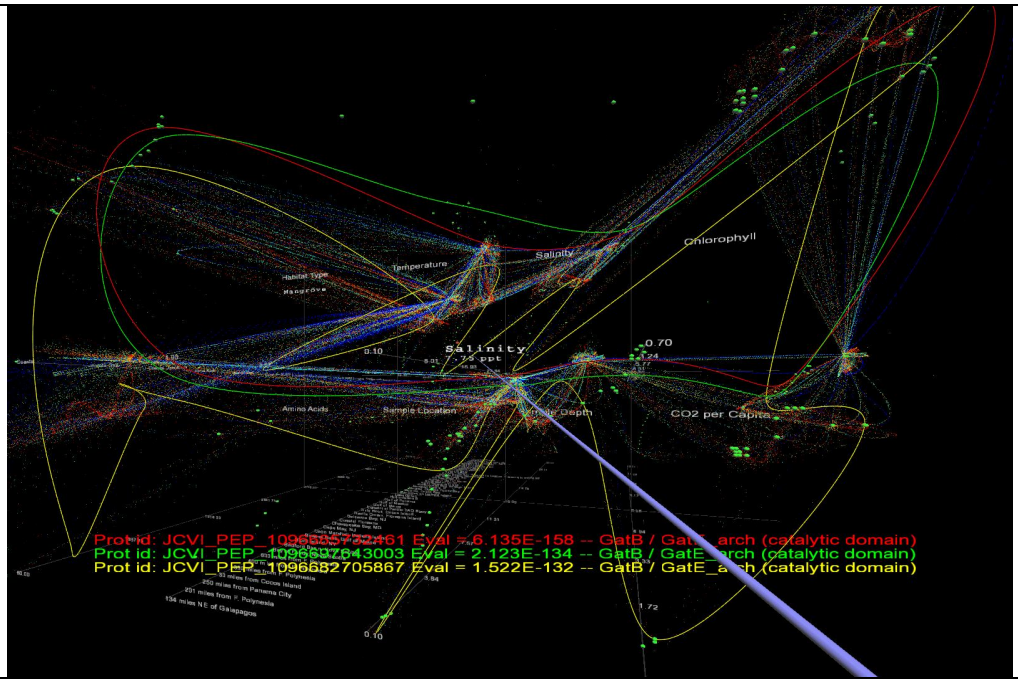

Fig. 11. BLAST results in MDE context: Each of the three colored lines (tracebacks) is a spatio-temporal signature encoding the set of metadata values associated with the top hits in a BLAST query result set. The top hit (red) and the second hit (green) each have a single set of metadata descriptors. Their values, while similar, are distinct and demonstrate that the sequences were identified at different sample locations. The third hit (yellow) is a sequence that has multiple sets of metadata descriptors (4 values for each category). It is a sequence that was sampled at a single physical location but algorithmically identified in collections of sequences from three additional sample sites. (Center) The purple colored pointer with associated "Salinity" value shows a user exploring the metadata space.

3. MAKING AND BREAKING PATTERNS

Metadata environments are an example of aesthetically impelled approaches for working with large and complex data in which user exploration makes and breaks patterns within a broader contextual scaffold. Our design of the MDE for the GOS metagenomics dataset piloted a method to integrate multiple data types to enable intuitive discovery-oriented browsing of large data collections by non-expert audiences. The prototype consists of a metadata driven virtual environment that is partitioned in to sub-regions, each representing specific metadata attributes used to describe the entire GOS data collection. This environment is instantiated within a computed fluid force. Each particle in the fluid corresponds to a single database record and movement reveals the specific values of all metadata descriptors for each record. Our future work envisions creating a generalized and scalable (extensible) MDE framework. We envision enabling users to construct environments based upon not only metadata descriptors for data sets, but any combination of variables, to create environments of arbitrary size and dimensionality and to dynamically adjust regions interactively at run time as a method for intuitive, open-ended exploration of very large data spaces. The design and implementation of MDEs are an example of dataremix-ing[18] in which data is conceptualized as akin to a fluid, or flow, with multiple state transitions rather than a raw material that is acted upon in discrete phases progressing along the data to knowledge pipeline. Our aesthetic and artistic approach to visualization and sonification of GOS data and contextual metadata presented here is data-driven yet non-hypothesis nor problem driven. It is a hybrid multi-scale strategy that merges quantitative and qualitative representation with the aim of supporting open-ended, discovery-oriented browsing and exploration of massive multidimensional data collections in ways that do not require a priori knowledge of the relationship between the underlying data and its mapping. Data is presented visually within the virtual environment as dynamic and abstract patterns in different positions relative to the overall virtual world coordinate system and the user (real-world) coordinates. Data is also presented not only as audio objects within the virtual environment, but using

spatialization strategies that position and move audio objects relative to the user according to both their interaction with the patterns, and the relation between and within data objects themselves. In a broader context, this work engages the concepts of "context' and "pattern" as framing and facilitating data exploration in circumstances where one may not know what one is looking for (e.g. detecting the unexpected, ideation, or hypothesis generation) and where the data can be accessed by a broad user base spanning researchers, citizen scientists, educators and the general public.

ACKNOWLEDGEMENTS

This work was supported in part by National Science Foundation IIS-0841031. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Additional support was provided by Calit2, CRCA, CAMERA, Ingenuity Festival, TimeLogic, CENS, NCMIR, EVL and SDSC. Installation hardware and software is provided by Da-Lite Screen Company, VRCO/Mechdyne, Meyer Sound, and mentalimages.

[1] Gantz J.F., Reinsel D. (2012) The Digital Universe in 2020:Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East– United States, IDC. Online: http://www.emc.com/leadership/digital-universe/iview/index.htm Last accessed: 8/31/13

[2] danah boyd & Kate Crawford (2012) Critical Questions For Big Data, Information, Communication & Society, 15:5, 662-679

[3] Ackoff, R.L. (1989) "From Data to Wisdom", Journal of Applied Systems Analysis, Volume 16, 1989 p 3-9

[4] Naur p. (1995) Knowing and the Mystique of Logic and Rules, Kluwer Academic

[5] Malina R.F., Strohecker S, LaFayette C, and Ione A. (2013) Steps to an Ecology of Networked Knowledge and Innovation: Enabling new forms of collaboration among sciences, engineering, arts, and design" http://seadnetwork.wordpress.com/draft-overview-of-a-report-on- the-sead-white-papers/ Last accessed: 8/31/13

[6] Cleveland, W, (1993) Visualizing Data, Hobart Press.

[7] Clark A. (2008) Supersizing the mind. Embodiment, Action and Cognitive Extension. Oxford University Press.

[8] ATLAS in silico, online at: http://atlasinsilico.net

[9] West R, Lewis JP, Margolis T, Schulze JP, Gossmann J, Tenedorio D, Singh R. (2009) Algorithmic Object As Natural Specimen: Meta Shape Grammar Objects From Atlas In Silico, Leonardo Electronic Almanac, Vol. 6, Issue 6 – 7, October 2009.

[10] J. Gossman, B. Hackbarth, and R. West, with T. Margolis, J.P. Lewis, and I. Mostafavi. (2008) Sclable Auditory Data Signatures for Discovery Oriented Browsing in an Expressive Context. Proceedings of the 14th International Conference on Auditory Display, June 24 - 28, 2008, Paris, France.

[11] West R, Gossmann J, Margolis T, Schulze JP, Lewis JP, Tenedorio D. (2009). Sensate abstraction: hybrid strategies for multi-dimensional data in expressive virtual reality contexts. Proceedings of the 21st Annual SPIE Symposium on Electronic Imaging, The Engineering Reality of Virtual Reality, 18-22 January 2009 San Jose, California, Volume 7238, pp. 72380I-72380I-11 (2009).

[12] Yooseph S, Sutton G, Rusch DB, Halpern AL, Williamson SJ, et al. "The Sorcerer II Global Ocean Sampling Expedition: Expanding the Universe of Protein Families." PLoS Biol 5(3) (2007).

[13] CAMERA (Community Cyberinfrastructure for Advanced Microbial Ecology Research and Analysis) portal: http://camera.calit2.net/

[14] Stam, J. (1999) Stable Fluids. Proc. SIGGRAPH 1999, ACM, 121-128.

[15] Lewis, J.P. , Rosenholtz, R., Fong, N., Neumann, U. (2004). VisualIDs: Automatic Distinctive Icons for Desktop Interfaces, ACM Trans. Graphics Volume 23, #3 (August 2004), 416-423.

[16] Altschul, S. F., et al. (1990) Basic Local Alignment Search Tool, Journal of Molecular Biology 215, 403-410.

[17] Altschul, S. F., et al., (1997) Gapped BLAST and PSI BLAST: A New Generation of Protein Database Search Programs," Nucleic Acids Research 25, No. 17, 3389-3402.

[18] West R., Malina R., Lewis J., Gresham-Lancaster S., Borsani A, Merlo B, Wang L. (2013) DataRemix: Designing the DataMade Through ArtScience Collaboration. In *Proceedings of the IEEE VIS Arts Program (VISAP)*, Atlanta, Georgia, October 2013.